1. A method for crawling a web site, the method comprising the steps of:

a) querying a web site server by a crawler program, wherein at least one page of the web site has a reference for executing by a browser to produce an address for a next page;

b) parsing such a reference from one of the web pages by the crawler program and sending the reference to an applet running in the browser; and

c) determining the address for the next page by the browser responsive to the reference and sending the address to the crawler.

2. The method of claim 2, the browser being configured to use a certain proxy, and refer to a resolver file for hostname-to-IP-address-resolution, and wherein the web site server has an IP address, the proxy for the browser has a certain IP address, and the resolver file indicates the certain IP address as the IP address for the web site server.

3. The method of claim 2, comprising the steps of:

adding an onload attribute to one of the web pages by the proxy;

defining an event handler for the onload attribute by the proxy, wherein the event handler sets a certain variable; and

polling the certain variable by the applet to determine when the page is loaded.

4. The method of claim 1, wherein the crawler is programmable to perform particular action sequences for generating the queries to the web server.

5. The method of claim 1, at least some of the web pages being dynamically generated by the server responsive to corresponding ones of the queries, comprising the step of:

processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server.

6. The method of claim 5, wherein at least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and wherein the step of processing the server generated web pages comprises the step of:

removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.

7. A method for reducing dynamic data generation on a web site server, the method comprising the steps of:

a) querying a web site server by a crawler program responsive to references from one web page to another in the web site, wherein the queries are for causing the server to generate web pages, at least some of the web pages being dynamically generated; and

b) processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server.

8. The method of claim 7, wherein at least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and wherein the step of processing the server generated web pages comprises the step of:

removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.

9. A computer program product for crawling a web site, the computer program product comprising:

a) first instructions for querying a web site server by a crawler program, wherein at least one page of the web site has a reference for executing by a browser to produce an address for a next page;

b) second instructions for parsing such a reference from one of the web pages by the crawler program and sending the reference to an applet running in the browser; and

c) third instructions for determining the address for the next page by the browser responsive to the reference and sending the address to the crawler.

10. The computer program product of claim 9, the browser being configured to use a certain proxy, and refer to a resolver file for hostname-to-IP-address-resolution, and wherein the web site server has an IP address, the proxy for the browser has a certain IP address, and the resolver file indicates the certain IP address as the IP address for the web site server.

11. The computer program product of claim 10, comprising:

fourth instructions for adding an onload attribute to one of the web pages by the proxy;

fifth instructions for defining an event handler for the onload attribute by the proxy, wherein the event handler sets a certain variable; and

sixth instructions for polling the certain variable by the applet to determine when the page is loaded.

12. The computer program product of claim 9, wherein the first instructions comprise instructions for causing the crawler to perform particular action sequences for generating the queries to the web server.

13. The computer program product of claim 1, at least some of the web pages being dynamically generated by the server responsive to corresponding ones of the queries, the computer program product comprising:

instructions for processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server.

14. The computer program product of claim 13, wherein at least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and wherein the instructions for processing the server generated web pages to generate corresponding processed versions of the web pages comprise:

instructions for removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.

15. A computer program product for reducing dynamic data generation on a web site server, the computer program product comprising:

first instructions for querying a web site server by a crawler program responsive to references from one web page to another in the web site, wherein the queries are for causing the server to generate web pages, at least some of the web pages being dynamically generated; and

second instructions for processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server.

16. The computer program product of claim 15, wherein at least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and wherein the seventh instructions comprise:

instructions for removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.

17. An apparatus for for crawling a web site, the apparatus comprising:

a processor connected a network,

a storage device connected to the processor and the network, wherein the storage device is for storing a program for controlling the processor, and wherein the processor is operative with the program to execute a crawler program and a browser program for performing the steps of:

a) querying a web site server by the crawler, wherein at least one page of the web site has a reference for executing by the browser to produce an address for a next page;

b) parsing such a reference from one of the web pages and sending the reference to an applet running in the browser; and

c) determining the address for the next page by the browser responsive to the reference and

sending the address to the crawler.

18. The apparatus of claim 17, the browser being configured to use a certain proxy, and refer to a resolver file for hostname-to-IP-address-resolution, and wherein the web site server has an IP address, the proxy for the browser has a certain IP address, and the resolver file indicates the certain IP address as the IP address for the web site server.

19. The apparatus of claim 18, wherein an onload attribute is added to one of the web pages by the proxy, and an event handler is defined for the onload attribute to set a certain variable, and wherein the processor is operative with the program for performing the step of:

polling the certain variable by the applet to determine when the page is loaded.

20. The apparatus of claim 17, wherein the processor is operative with the program for causing the crawler to perform particular action sequences for generating the queries to the web server.

21. The apparatus of claim 17, at least some of the web pages being dynamically generated by the server responsive to corresponding ones of the queries, wherein the processor is operative with the program for performing the step of:

processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server.

22. The apparatus of claim 21, wherein at least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and wherein the step of processing the server generated web pages comprises the step of:

removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.

23. An apparatus for reducing dynamic data generation on a web site server, the apparatus comprising:

a processor connected a network,

a storage device connected to the processor and the network, wherein the storage device is for storing a program for controlling the processor, and wherein the processor is operative with the program to execute a crawler program and a browser program for performing the steps of:

a) querying a web site server by the crawler responsive to references from one web page to another in the web site, wherein the queries are for causing the server to generate web pages, at least some of the web pages being dynamically generated; and

b) processing the server generated web pages to generate corresponding processed versions of the web pages, so that the processed versions can be served in response to future queries, reducing dynamic generation of web pages by the server.

24. The apparatus of claim 23, wherein at least a first such server generated web page has included in it an operation that would cause the server to dynamically generate a second web page if the first page were used to generate further requests to the server, and wherein the step of processing the server generated web pages comprises the step of:

removing the operation from the first server generated web page and replacing the operation with a reference to a version of another of the server generated web pages.